SEQUENCE LISTING

<110> SHOWA DENKO K.K.

<120> NOVEL RHODOCOCCUS BACTERIA, NITRILASE GENE, NITRYL HYDRATASE GENE AND AMIDASE GENE FROM RHODOCOCCUS BACTERIUM, AND PROCESS FOR PRODUCING CARBOXYLIC ACIDS USING THEM

<130> Q64574

<140> 09/869,142
<141> 2001-06-26

<150> USSN 60/183,754
<151> 2000-02-22

<150> USSN 60/183,821
<151> 2000-02-22

<150> JPA 2000-107855
<151> 2000-04-10

<150> JPA 2000-021797
<151> 2000-01-26

<150> JPA 11-303212
<151> 1999-10-26

<160> 7

<170> PatentIn version 3.1

<210> 1
<211> 1531
<212> DNA
<213> Rhodococcus sp.

<220>
<221> exon
<222> (324)..(1421)
<223>

<400> 1
agcttgacca tgattacgaa ttcgagctcg gtacccgggg atcgaaccag caacggggac 60

gcacagtcga cgtagacctc gacctatccg ccgttccgca gaaggacacc gaccaccacc 120

acttcaacat ccttcaacgt gcccggccag tccttcgacg aatcgaaacg gcgaagagcc 180

gcctcggacc ccccggccga accgctcgat gaactcccct acacgggtgg cgcagaatgc 240

caggacccgt gtcattccac gtcaattcac gcgccttttc acctcgtact gtcctgccaa 300

acacaagcaa cggaggtacg gac atg gtc gaa tac aca aac aca ttc aaa gtt 353
Met Val Glu Tyr Thr Asn Thr Phe Lys Val
1 5 10

gct gcg gtg cag gca cag cct gtg tgg ttc gac gcg gcc aaa acg gtc 401
Ala Ala Val Gln Ala Gln Pro Val Trp Phe Asp Ala Ala Lys Thr Val

```
                15                  20                  25

gac aag acc gtg tcc atc atc gcg gaa gca gcc cgg aac ggg tgc gag      449
Asp Lys Thr Val Ser Ile Ile Ala Glu Ala Ala Arg Asn Gly Cys Glu
            30                  35                  40

ctc gtt gcg ttt ccc gag gta ttc atc ccg ggg tac ccg tac cac atc      497
Leu Val Ala Phe Pro Glu Val Phe Ile Pro Gly Tyr Pro Tyr His Ile
        45                  50                  55

tgg gtc gac agc ccg ctc gcc gga atg gcg aag ttc gcc gtg cgc tac      545
Trp Val Asp Ser Pro Leu Ala Gly Met Ala Lys Phe Ala Val Arg Tyr
    60                  65                  70

cac gag aat tcc ctg acg atg gac agc ccg cac gta cag cgg ttg ctc      593
His Glu Asn Ser Leu Thr Met Asp Ser Pro His Val Gln Arg Leu Leu
75                  80                  85                  90

gat gcc gcc cgc gac cac aac atc gcc gta gtg gtg gga atc agc gag      641
Asp Ala Ala Arg Asp His Asn Ile Ala Val Val Val Gly Ile Ser Glu
                95                  100                 105

cgg gat ggc ggc agc ttg tac atg acc cag ctc atc atc gac gcc gat      689
Arg Asp Gly Gly Ser Leu Tyr Met Thr Gln Leu Ile Ile Asp Ala Asp
            110                 115                 120

ggg caa ctg gtc gcc cga cgc cgc aag ctc aag ccc acc cac gtc gag      737
Gly Gln Leu Val Ala Arg Arg Arg Lys Leu Lys Pro Thr His Val Glu
        125                 130                 135

cgt tcg gta tac gga gaa gga aac ggc tcg gat atc tcc gtg tac gac      785
Arg Ser Val Tyr Gly Glu Gly Asn Gly Ser Asp Ile Ser Val Tyr Asp
    140                 145                 150

atg cct ttc gca cgg ctt ggc gcg ctc aac tgc tgg gag cat ttc cag      833
Met Pro Phe Ala Arg Leu Gly Ala Leu Asn Cys Trp Glu His Phe Gln
155                 160                 165                 170

acg ctc acc aag tac gca atg tac tcg atg cac gag cag gtg cac gtc      881
Thr Leu Thr Lys Tyr Ala Met Tyr Ser Met His Glu Gln Val His Val
                175                 180                 185

gcg agc tgg cct ggc atg tcg ctg tac cag ccg gag gtc ccc gca ttc      929
Ala Ser Trp Pro Gly Met Ser Leu Tyr Gln Pro Glu Val Pro Ala Phe
            190                 195                 200

ggt gtc gat gcc cag ctc acg gcc acg cgt atg tac gca ctc gag gga      977
Gly Val Asp Ala Gln Leu Thr Ala Thr Arg Met Tyr Ala Leu Glu Gly
        205                 210                 215

caa acc ttc gtg gtc tgc acc acc cag gtg gtc aca ccg gag gcc cac     1025
Gln Thr Phe Val Val Cys Thr Thr Gln Val Val Thr Pro Glu Ala His
    220                 225                 230

gag ttc ttc tgc gag aac gag gaa cag cga atg ttg atc ggc cga ggc     1073
Glu Phe Phe Cys Glu Asn Glu Glu Gln Arg Met Leu Ile Gly Arg Gly
235                 240                 245                 250

gga ggt ttc gcg cgc atc atc ggg ccc gac ggc cgc gat ctc gca act     1121
Gly Gly Phe Ala Arg Ile Ile Gly Pro Asp Gly Arg Asp Leu Ala Thr
```

```
                255                 260                 265

cct ctc gcc gaa gat gag gag ggg atc ctc tac gcc gac atc gat ctg     1169
Pro Leu Ala Glu Asp Glu Glu Gly Ile Leu Tyr Ala Asp Ile Asp Leu
            270                 275                 280

tct gcg atc acc ttg gcg aag cag gcc gct gac ccc gtg ggc cac tac     1217
Ser Ala Ile Thr Leu Ala Lys Gln Ala Ala Asp Pro Val Gly His Tyr
        285                 290                 295

tca cgg ccg gat gtg ctg tcg ctg aac ttc aac cag cgc cgc acc acg     1265
Ser Arg Pro Asp Val Leu Ser Leu Asn Phe Asn Gln Arg Arg Thr Thr
    300                 305                 310

ccc gtc aac acc cca ctt tcc acc atc cat gcc acg cac acg ttc gtg     1313
Pro Val Asn Thr Pro Leu Ser Thr Ile His Ala Thr His Thr Phe Val
315                 320                 325                 330

ccg cag ttc ggg gca ctc gac ggc gtc cgt gag ctc aac gga gcg gac     1361
Pro Gln Phe Gly Ala Leu Asp Gly Val Arg Glu Leu Asn Gly Ala Asp
                335                 340                 345

gaa cag cgc gca ttg ccc tcc aca cat tcc gac gag acg gac cgg gcg     1409
Glu Gln Arg Ala Leu Pro Ser Thr His Ser Asp Glu Thr Asp Arg Ala
            350                 355                 360

aca gcc acc ctc tgactcgggc gcacccgtgg cgcctccgaa gcgccacggg         1461
Thr Ala Thr Leu
        365

tgtgtgaagg ggcgagacag gggaatcgga ggatccccgg gtaccgagct cgaattcgta   1521

atcatggtca                                                         1531
```

<210> 2
<211> 366
<212> PRT
<213> Rhodococcus sp.

<400> 2

```
Met Val Glu Tyr Thr Asn Thr Phe Lys Val Ala Ala Val Gln Ala Gln
1               5                   10                  15

Pro Val Trp Phe Asp Ala Ala Lys Thr Val Asp Lys Thr Val Ser Ile
            20                  25                  30

Ile Ala Glu Ala Ala Arg Asn Gly Cys Glu Leu Val Ala Phe Pro Glu
        35                  40                  45

Val Phe Ile Pro Gly Tyr Pro Tyr His Ile Trp Val Asp Ser Pro Leu
    50                  55                  60

Ala Gly Met Ala Lys Phe Ala Val Arg Tyr His Glu Asn Ser Leu Thr
65                  70                  75                  80
```

Met Asp Ser Pro His Val Gln Arg Leu Leu Asp Ala Ala Arg Asp His
85 90 95

Asn Ile Ala Val Val Val Gly Ile Ser Glu Arg Asp Gly Gly Ser Leu
100 105 110

Tyr Met Thr Gln Leu Ile Ile Asp Ala Asp Gly Gln Leu Val Ala Arg
115 120 125

Arg Arg Lys Leu Lys Pro Thr His Val Glu Arg Ser Val Tyr Gly Glu
130 135 140

Gly Asn Gly Ser Asp Ile Ser Val Tyr Asp Met Pro Phe Ala Arg Leu
145 150 155 160

Gly Ala Leu Asn Cys Trp Glu His Phe Gln Thr Leu Thr Lys Tyr Ala
165 170 175

Met Tyr Ser Met His Glu Gln Val His Val Ala Ser Trp Pro Gly Met
180 185 190

Ser Leu Tyr Gln Pro Glu Val Pro Ala Phe Gly Val Asp Ala Gln Leu
195 200 205

Thr Ala Thr Arg Met Tyr Ala Leu Glu Gly Gln Thr Phe Val Val Cys
210 215 220

Thr Thr Gln Val Val Thr Pro Glu Ala His Glu Phe Phe Cys Glu Asn
225 230 235 240

Glu Glu Gln Arg Met Leu Ile Gly Arg Gly Gly Gly Phe Ala Arg Ile
245 250 255

Ile Gly Pro Asp Gly Arg Asp Leu Ala Thr Pro Leu Ala Glu Asp Glu
260 265 270

Glu Gly Ile Leu Tyr Ala Asp Ile Asp Leu Ser Ala Ile Thr Leu Ala
275 280 285

Lys Gln Ala Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu
290 295 300

Ser Leu Asn Phe Asn Gln Arg Arg Thr Thr Pro Val Asn Thr Pro Leu
305 310 315 320

Ser Thr Ile His Ala Thr His Thr Phe Val Pro Gln Phe Gly Ala Leu
325 330 335

Asp Gly Val Arg Glu Leu Asn Gly Ala Asp Glu Gln Arg Ala Leu Pro
340 345 350

Ser Thr His Ser Asp Glu Thr Asp Arg Ala Thr Ala Thr Leu
355 360 365

<210> 3
<211> 2822
<212> DNA
<213> Rhodococcus sp.

<220>
<221> CDS
<222> (1379)..(2068)
<223> nitrile hydratase beta subunit

<220>
<221> CDS
<222> (2082)..(2693)
<223> nitrile hydratase alpha subunit

<400> 3
ctagaggatc tcggtcatcg cgataccatc gttgcggacg atgatgtcca atacgtacca 60

ctggtccgcg gtcaacttct cttgatcgac cacgttatgg attctacgac tcagggaccg 120

gctcacggct tccagggcgc ctccgaccaa aggtgatcga acgacatttc cggattcagc 180

caccgcttcc gactcgatca ttcctgtccc tccccgtcca cgcgcagttg atcttacctc 240

ctcatcaaga ggatatccac tgaacgaatt atttcaagtg gaagtacttg gagtcgatcc 300

tacacgtgag tggacgatgc ctgggcgcta gtcggatgtg caacccaccc acccccctcct 360

cccgcctacg ccgaagaccg gaaccggcgt cgtccctgcc tgccgtctct ggcaactgtt 420

gtgaacgccc gagcggccct cacggctctt cagttggcgc ggatcgccat ggcggacgtc 480

gcccacggcg ggacctacgc atcttcggcc ggaaggcagc cgcggtcacg aacacctagc 540

ggcagtcgag cacctgagac gaaggccgcc ggcgtcctgt cccggaaatc cgcagcccag 600

ccgtgacagc caacagtcgt ggcggttccc tcccctcctа gggtctttga ctcggcgcca 660

acgcctgcga gggcgctcgt cgcggaccac ttgtcgaggt cggtgccgca cgtcaccgag 720

cgcacccttc ttcgtgctct gcgcatcggc ccggaccgcg accgcggcaa cactacgacg 780

tctgacaatg ctgatcccct gccgccgccg ttggacgacc acagttgcta cgagcatgcg 840

```
gagccaacca taggcatcat gcgatcgccg gagtcttcat cctattttgg gatgcgcagg    900

attaacacat ctacacattg acatccgttc cgatgtgaag taaaaattgt cacgtagggc    960

ggcaggcgaa gtctgcagct cgaacatcga agggtgggag ccgagagatc ggagacgcag   1020

acacccggag ggaacttagc ctcccggacc gatgcgtgtc ctggcaacgc ctcaagattc   1080

agcgcaagcg attcaatctt gttacttcca gaaccgaatc acgtccccgt agtgtgcggg   1140

gagagcgccc gaacgcaggg atggtatcca tgcgcccctt ctcttttcga acgagaaccg   1200

gccggtacag tcaatccgga cacattgtga cgccgttcaa cgattgttgt gctgtgaagg   1260

attcactcaa gccaactgat atcgccattc cgttgccgga acatttgacg ccttctccct   1320

acgagtagaa gccagctgga ccctctttga gcccagctcc gatgaaagga atgaggaa      1378

atg gat ggt atc cac gac aca ggc ggc atg acc gga tac gga ccg gtc      1426
Met Asp Gly Ile His Asp Thr Gly Gly Met Thr Gly Tyr Gly Pro Val
1               5                  10                  15

ccc tat cag aag gac gag ccc ttc ttc cac tac gag tgg gag ggt cga      1474
Pro Tyr Gln Lys Asp Glu Pro Phe Phe His Tyr Glu Trp Glu Gly Arg
            20                  25                  30

acc ctg tcg att ctg acc tgg atg cat ctc aag ggc atg tcg tgg tgg      1522
Thr Leu Ser Ile Leu Thr Trp Met His Leu Lys Gly Met Ser Trp Trp
        35                  40                  45

gac aag tcg cgg ttc ttc cgg gag tcg atg ggg aac gaa aac tac gtc      1570
Asp Lys Ser Arg Phe Phe Arg Glu Ser Met Gly Asn Glu Asn Tyr Val
    50                  55                  60

aac gag att cgc aac tcg tac tac acc cac tgg ctg agt gcg gcg gaa      1618
Asn Glu Ile Arg Asn Ser Tyr Tyr Thr His Trp Leu Ser Ala Ala Glu
65                  70                  75                  80

cgt atc ctc gtc gcc gac aag atc atc acc gaa gaa gag cga aag cac      1666
Arg Ile Leu Val Ala Asp Lys Ile Ile Thr Glu Glu Glu Arg Lys His
                85                  90                  95

cgc gtg cag gag atc ctc gag ggt cgg tac acg gac agg aac ccg tcg      1714
Arg Val Gln Glu Ile Leu Glu Gly Arg Tyr Thr Asp Arg Asn Pro Ser
            100                 105                 110

cgg aag ttc gat ccg gcc gag atc gag aag gcg atc gag agg ctt cac      1762
Arg Lys Phe Asp Pro Ala Glu Ile Glu Lys Ala Ile Glu Arg Leu His
        115                 120                 125

gag ccc cac tcc cta gtg ctt cca gga gcg gag ccg agt ttc tcc ctc      1810
Glu Pro His Ser Leu Val Leu Pro Gly Ala Glu Pro Ser Phe Ser Leu
    130                 135                 140

ggt gac aag gtc aaa gtg aag aac atg aac ccg ctg gga cac aca cgg      1858
Gly Asp Lys Val Lys Val Lys Asn Met Asn Pro Leu Gly His Thr Arg
145                 150                 155                 160

tgc ccg aag tat gtg cgg aac aga atc ggg gaa atc gtc acc tcc cac      1906
Cys Pro Lys Tyr Val Arg Asn Arg Ile Gly Glu Ile Val Thr Ser His
```

```
                165                 170                 175

ggg tgc cag atc tat ccc gag agc agc tcc gcc ggc ctc ggc gac gat      1954
Gly Cys Gln Ile Tyr Pro Glu Ser Ser Ser Ala Gly Leu Gly Asp Asp
            180                 185                 190

ccc cgc ccg ctc tac acg gtc gcg ttt tcc gcc cag gaa ctg tgg ggc      2002
Pro Arg Pro Leu Tyr Thr Val Ala Phe Ser Ala Gln Glu Leu Trp Gly
        195                 200                 205

gac gac gga aac ggg aaa gac gta gtg tgc gtc gat ctc tgg gaa ccg      2050
Asp Asp Gly Asn Gly Lys Asp Val Val Cys Val Asp Leu Trp Glu Pro
    210                 215                 220

tac ctg atc tct gcg tga aaggaatacg ata gtg agc gag cac gtc aat      2099
Tyr Leu Ile Ser Ala                 Val Ser Glu His Val Asn
225                                 230                 235

aag tac acg gag tac gag gca cgt acc aag gca atc gaa acc ttg ctg      2147
Lys Tyr Thr Glu Tyr Glu Ala Arg Thr Lys Ala Ile Glu Thr Leu Leu
                240                 245                 250

tac gag cga ggg ctc atc acg ccc gcc gcg gtc gac cga gtc gtt tcg      2195
Tyr Glu Arg Gly Leu Ile Thr Pro Ala Ala Val Asp Arg Val Val Ser
            255                 260                 265

tac tac gag aac gag atc ggc ccg atg ggc ggt gcc aag gtc gtg gcc      2243
Tyr Tyr Glu Asn Glu Ile Gly Pro Met Gly Gly Ala Lys Val Val Ala
        270                 275                 280

aag tcc tgg gtg gac cct gag tac cgc aag tgg ctc gaa gaa gac gcg      2291
Lys Ser Trp Val Asp Pro Glu Tyr Arg Lys Trp Leu Glu Glu Asp Ala
    285                 290                 295

acg gcc gcg atg gcg tca ttg ggc tat gcc ggc gag cag gca cac cag      2339
Thr Ala Ala Met Ala Ser Leu Gly Tyr Ala Gly Glu Gln Ala His Gln
300                 305                 310                 315

atc tcg gcc gtc ttc aac gac tcc caa aca cat cac gta gtg gtg tgc      2387
Ile Ser Ala Val Phe Asn Asp Ser Gln Thr His His Val Val Val Cys
                320                 325                 330

act ctg tgt tcg tgc tat ccg tgg ccg gtg ctt ggc ctc ccg ccc gcc      2435
Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val Leu Gly Leu Pro Pro Ala
            335                 340                 345

tgg tac aag agc atg gag tac cgg tcc cga gtg gta gca gac cct cgt      2483
Trp Tyr Lys Ser Met Glu Tyr Arg Ser Arg Val Val Ala Asp Pro Arg
        350                 355                 360

gga gta ctc aag cgc gat ttc ggg ttc gac atc ccc gat gag gtg gag      2531
Gly Val Leu Lys Arg Asp Phe Gly Phe Asp Ile Pro Asp Glu Val Glu
    365                 370                 375

gtc agg gtt tgg gac agc agc tcc gaa atc cgc tac atc gtc atc ccg      2579
Val Arg Val Trp Asp Ser Ser Ser Glu Ile Arg Tyr Ile Val Ile Pro
380                 385                 390                 395

gaa cgg ccg gcc ggc acc gac ggt tgg tcc gag gac gag ctg gcg aag      2627
Glu Arg Pro Ala Gly Thr Asp Gly Trp Ser Glu Asp Glu Leu Ala Lys
```

```
                400                 405                 410

ctg gtg agt cgg gac tcg atg atc ggt gtc agt aat gcg ctc aca ccg     2675
Leu Val Ser Arg Asp Ser Met Ile Gly Val Ser Asn Ala Leu Thr Pro
            415                 420                 425

cag gaa gtg atc gta tga gtgaagacac actcactgat cggctcccgg           2723
Gln Glu Val Ile Val
        430

cgactgggac cgccgcaccg ccccgcgaca atggcgagct tgtattcacc gagccttggg   2783

aagcaacggc attcggggtc gccatcgcgc tttcggatc                         2822


<210>  4
<211>  229
<212>  PRT
<213>  Rhodococcus sp.

<400>  4

Met Asp Gly Ile His Asp Thr Gly Gly Met Thr Gly Tyr Gly Pro Val
1               5                   10                  15

Pro Tyr Gln Lys Asp Glu Pro Phe Phe His Tyr Glu Trp Glu Gly Arg
            20                  25                  30

Thr Leu Ser Ile Leu Thr Trp Met His Leu Lys Gly Met Ser Trp Trp
        35                  40                  45

Asp Lys Ser Arg Phe Phe Arg Glu Ser Met Gly Asn Glu Asn Tyr Val
    50                  55                  60

Asn Glu Ile Arg Asn Ser Tyr Tyr Thr His Trp Leu Ser Ala Ala Glu
65                  70                  75                  80

Arg Ile Leu Val Ala Asp Lys Ile Ile Thr Glu Glu Glu Arg Lys His
                85                  90                  95

Arg Val Gln Glu Ile Leu Glu Gly Arg Tyr Thr Asp Arg Asn Pro Ser
            100                 105                 110

Arg Lys Phe Asp Pro Ala Glu Ile Glu Lys Ala Ile Glu Arg Leu His
        115                 120                 125

Glu Pro His Ser Leu Val Leu Pro Gly Ala Glu Pro Ser Phe Ser Leu
    130                 135                 140

Gly Asp Lys Val Lys Val Lys Asn Met Asn Pro Leu Gly His Thr Arg
145                 150                 155                 160
```

```
Cys Pro Lys Tyr Val Arg Asn Arg Ile Gly Glu Ile Val Thr Ser His
                165                 170                 175

Gly Cys Gln Ile Tyr Pro Glu Ser Ser Ser Ala Gly Leu Gly Asp Asp
            180                 185                 190

Pro Arg Pro Leu Tyr Thr Val Ala Phe Ser Ala Gln Glu Leu Trp Gly
        195                 200                 205

Asp Asp Gly Asn Gly Lys Asp Val Val Cys Val Asp Leu Trp Glu Pro
    210                 215                 220

Tyr Leu Ile Ser Ala
225
```

<210> 5
<211> 203
<212> PRT
<213> Rhodococcus sp.

<400> 5

```
Val Ser Glu His Val Asn Lys Tyr Thr Glu Tyr Glu Ala Arg Thr Lys
1               5                   10                  15

Ala Ile Glu Thr Leu Leu Tyr Glu Arg Gly Leu Ile Thr Pro Ala Ala
            20                  25                  30

Val Asp Arg Val Val Ser Tyr Tyr Glu Asn Glu Ile Gly Pro Met Gly
        35                  40                  45

Gly Ala Lys Val Val Ala Lys Ser Trp Val Asp Pro Glu Tyr Arg Lys
    50                  55                  60

Trp Leu Glu Glu Asp Ala Thr Ala Ala Met Ala Ser Leu Gly Tyr Ala
65                  70                  75                  80

Gly Glu Gln Ala His Gln Ile Ser Ala Val Phe Asn Asp Ser Gln Thr
                85                  90                  95

His His Val Val Val Cys Thr Leu Cys Ser Cys Tyr Pro Trp Pro Val
            100                 105                 110

Leu Gly Leu Pro Pro Ala Trp Tyr Lys Ser Met Glu Tyr Arg Ser Arg
        115                 120                 125
```

```
Val Val Ala Asp Pro Arg Gly Val Leu Lys Arg Asp Phe Gly Phe Asp
    130                 135                 140

Ile Pro Asp Glu Val Glu Val Arg Val Trp Asp Ser Ser Ser Glu Ile
145                 150                 155                 160

Arg Tyr Ile Val Ile Pro Glu Arg Pro Ala Gly Thr Asp Gly Trp Ser
                165                 170                 175

Glu Asp Glu Leu Ala Lys Leu Val Ser Arg Asp Ser Met Ile Gly Val
            180                 185                 190

Ser Asn Ala Leu Thr Pro Gln Glu Val Ile Val
        195                 200
```

<210> 6
<211> 2822
<212> DNA
<213> Rhodococcus sp.

<220>
<221> misc_feature
<222> (556)..(556)
<223> "n" may be a, c, g or t

<220>
<221> misc_feature
<222> (592)..(592)
<223> "n" may be a, c, g or t

<220>
<221> misc_feature
<222> (896)..(896)
<223> "n" may be a, c, g or t

<220>
<221> CDS
<222> (1094)..(2491)
<223> amidase

<400> 6
tgattacgaa ttcgagctcg gtacccgggg atcacttcgg ccagagggtg acggcgaaat 60

cgggcctcga tctccgcgtc cacggcgttg atacgtgtgt cgaggtcgat caccgcctgc 120

gccaattcgg cgaccagttc ggcagcgaca tcttcccccg gcaaccgcac ggtctgcgcc 180

ttcgcggcgg tgactgcggc ccgggcgatc gattcggcgt ggcgcacccc ggccccggtg 240

agcattgcgg ccagtcgggc cgccccgacg cggcggatcg ctttcggtcg ctggtagcgg 300

```
gccagcagca ccacccagcc ccggtccgag gagatctgcg cgacgcgttc gagtccgggg    360

cagatcgcga cgagttgctg acgcagccgg ttgatggtcc gggtacggtc ggcgaccaga    420

tcggtgcggt ggccggtgag catctgcagc tcacggatca actcgtcgtc gggacgcaga    480

acgggcaggt ccgaccgcat ccgggactga tcggcgatca cccgggcgtc gcgggcgtcg    540

gtcttggctt cgccgncgcg gtagaccgac gatgcctgcc acaccgaacg tncggacagg    600

tagcgcaccg gtttcccggc gtcggccagc acagtcagca acaaggtgac gtaggcggtg    660

gtcagatcca ccgtccacga caccgtctcg gtgagtgcgt cgatctcggt gatcaccgca    720

cggatcgttg cttcgtcgtt gtcacatcgc cgcgacagca ccaccgtccc ggaggtgtcg    780

agtacgcata tccagtggtg ttctttgccg acgtcgactc ctgcccacag ttgcgaaccg    840

gtcatcggat ttcctcgttt tcgcttgtgt tccggcctgg ccccgatgga cgcctncgcc    900

ggcatttcct taaacaagcg atcatgcgca gatctcaatc agcggtccag aggtgtccag    960

acaggtcggg tggccagtcc tttcaagccc cactcgagag tgggcaaacc ttatgcagcc   1020

tcgccggcct gcccgggtta cagctcaacg taactctcac gaagtaactg cacctacgaa   1080

cttaaggaac ctc atg tct tcg ttg act ccc ccc aat tcc aac caa atg        1129
               Met Ser Ser Leu Thr Pro Pro Asn Ser Asn Gln Met
               1               5                   10

tcg gcc ctg aac aac cac ttc cga ttc gga ctg acg acg ccg gaa ctc      1177
Ser Ala Leu Asn Asn His Phe Arg Phe Gly Leu Thr Thr Pro Glu Leu
        15                  20                  25

gaa gag ttc gca ccg gcc ctc gaa gcg acg ctc gcg tcc tcc gaa acc      1225
Glu Glu Phe Ala Pro Ala Leu Glu Ala Thr Leu Ala Ser Ser Glu Thr
    30                  35                  40

gtc gaa cgc ctc tac gag cgc acc gcg ccc gag ccg cct cag cgg tca      1273
Val Glu Arg Leu Tyr Glu Arg Thr Ala Pro Glu Pro Pro Gln Arg Ser
45                  50                  55                  60

tgg acc tca ccc acg gcg gac gag aac ccg ctg agc gcc tgg tac gtc      1321
Trp Thr Ser Pro Thr Ala Asp Glu Asn Pro Leu Ser Ala Trp Tyr Val
                65                  70                  75

acc acc tcg atc agc gaa acc gac gaa ggc ccc ctc gcc ggg cga acg      1369
Thr Thr Ser Ile Ser Glu Thr Asp Glu Gly Pro Leu Ala Gly Arg Thr
            80                  85                  90

gtc gcc gtg aaa gac aac gtc gca gtc gcc ggc gtg ccg atg atg aac      1417
Val Ala Val Lys Asp Asn Val Ala Val Ala Gly Val Pro Met Met Asn
        95                  100                 105

ggc tcc cga acc gtc gag ggc ttc acc ccc cgc tac gac gcc acc gtc      1465
Gly Ser Arg Thr Val Glu Gly Phe Thr Pro Arg Tyr Asp Ala Thr Val
    110                 115                 120

gta cgc cga ctg ctc gac gcc ggc gca acc atc acc ggc aaa gcg gtg      1513
```

```
Val Arg Arg Leu Leu Asp Ala Gly Ala Thr Ile Thr Gly Lys Ala Val
125                 130                 135                 140

tgc gaa gat ctc tgc ttc tcc ggc gcc agc ttc act tcc cac ccc cag      1561
Cys Glu Asp Leu Cys Phe Ser Gly Ala Ser Phe Thr Ser His Pro Gln
                145                 150                 155

ccg gtc cgc aac ccc tgg gac gaa agc cgc atc acc ggc ggc tcg tcc      1609
Pro Val Arg Asn Pro Trp Asp Glu Ser Arg Ile Thr Gly Gly Ser Ser
            160                 165                 170

agc ggc agc ggc gcc ctg gtc gcc agc ggc cag gtg gat atg gca gtc      1657
Ser Gly Ser Gly Ala Leu Val Ala Ser Gly Gln Val Asp Met Ala Val
        175                 180                 185

ggc ggc gac cag ggc ggt tcg atc cgc atc ccc gcc gcg ttc tgc ggc      1705
Gly Gly Asp Gln Gly Gly Ser Ile Arg Ile Pro Ala Ala Phe Cys Gly
    190                 195                 200

atc gtc gga cac aaa ccc acc cac gga ctg gtc ccc tat acg gga gca      1753
Ile Val Gly His Lys Pro Thr His Gly Leu Val Pro Tyr Thr Gly Ala
205                 210                 215                 220

ttt ccc atc gaa cga acc atc gac cac ctc ggt ccg atg acg cgc acg      1801
Phe Pro Ile Glu Arg Thr Ile Asp His Leu Gly Pro Met Thr Arg Thr
                225                 230                 235

gtc agc gac gcc gcc gca atg ctc acc gtc ctc gcc ggc acc gac ggc      1849
Val Ser Asp Ala Ala Ala Met Leu Thr Val Leu Ala Gly Thr Asp Gly
            240                 245                 250

ctc gat ccc cga cag acc cac cgg atc gaa ccg gtg gac tac ctc gcg      1897
Leu Asp Pro Arg Gln Thr His Arg Ile Glu Pro Val Asp Tyr Leu Ala
        255                 260                 265

gcg ctg gcc gaa ccc gca tcg ggt ctg cgc gtg ggt gtg gtc acc gaa      1945
Ala Leu Ala Glu Pro Ala Ser Gly Leu Arg Val Gly Val Val Thr Glu
    270                 275                 280

ggc ttc gac acc cct gtc tcc gac gct gcc gtc gac aat gcc gtg cgc      1993
Gly Phe Asp Thr Pro Val Ser Asp Ala Ala Val Asp Asn Ala Val Arg
285                 290                 295                 300

acc gcc atc ggc gta ctg cgc tcg gcc gga ctt acc gtc gaa gag gtc      2041
Thr Ala Ile Gly Val Leu Arg Ser Ala Gly Leu Thr Val Glu Glu Val
                305                 310                 315

tcg atc ccc tgg cac ctc gat gcg atg gcc gtc tgg aac gtg atc gac      2089
Ser Ile Pro Trp His Leu Asp Ala Met Ala Val Trp Asn Val Ile Asp
            320                 325                 330

cgg gcc gac gac gaa ttc gaa gcc ttc ctg ctg cag gtg ctc gac gag      2137
Arg Ala Asp Asp Glu Phe Glu Ala Phe Leu Leu Gln Val Leu Asp Glu
        335                 340                 345

aac gcc gtc acc atc ccc gaa ctc gga cag gtg cgg gcg cag acg ccg      2185
Asn Ala Val Thr Ile Pro Glu Leu Gly Gln Val Arg Ala Gln Thr Pro
    350                 355                 360

cgc tcg tgg tgc tca cct cga acc gca ccc gcg agg tgc acg acg ccc      2233
```

```
Arg Ser Trp Cys Ser Pro Arg Thr Ala Pro Ala Arg Cys Thr Thr Pro
365                 370                 375                 380

tca aac gcc gct gcc tgt acc act ggc tcg aac acc ccg acc tcg cgc     2281
Ser Asn Ala Ala Ala Cys Thr Thr Gly Ser Asn Thr Pro Thr Ser Arg
                385                 390                 395

ggg aag tgg aga tcc tgc gcc gcc gca tcc cgg gca tcg acg aac acc     2329
Gly Lys Trp Arg Ser Cys Ala Ala Ala Ser Arg Ala Ser Thr Asn Thr
            400                 405                 410

tcg cgg cgc agg tcg ccc acg ccg tgc agg cca tgc gcg gga tgg acc     2377
Ser Arg Arg Arg Ser Pro Thr Pro Cys Arg Pro Cys Ala Gly Trp Thr
        415                 420                 425

tgc tca aac cac ccg ggg tcg cgg agt cgc tgg act ggg cac gag cgc     2425
Cys Ser Asn His Pro Gly Ser Arg Ser Arg Trp Thr Gly His Glu Arg
    430                 435                 440

tgc ggg aac tcg acc gcg acg tgc tcg acg cga cga ccg cgg ccg cga     2473
Cys Gly Asn Ser Thr Ala Thr Cys Ser Thr Arg Arg Pro Arg Pro Arg
445                 450                 455                 460

ccc tcg gtg ccg tcc tga agtaccggga ggacctcgac cgagtggtcc            2521
Pro Ser Val Pro Ser
                465

gcaccgggct cgaccggctc ctgacggggt gacagcggcg atgacgacga ccaccgacgc   2581

cgggggttcc ctcgtcggac tcaccggctt cacccgcgcc ctcgccgcgg ccggcctgtc   2641

cgtcgcctcg gacgccaccg tggcctacct gcgcgccctg cgcgagatcg acctcggcga   2701

ccgccgtcag gtgtactggg ccgggcgcgc caccctgtgc cacgaccccg acgacatccc   2761

ccgctacgac ctcgcgttcg agagctggtt cggcggaacg gcacccgacg tgacgtcgcc   2821

g                                                                   2822
```

<210> 7
<211> 465
<212> PRT
<213> Rhodococcus sp.

<220>
<221> misc_feature
<222> (556)..(556)
<223> "n" may be a, c, g or t

<220>
<221> misc_feature
<222> (592)..(592)
<223> "n" may be a, c, g or t

<220>
<221> misc_feature
<222> (896)..(896)
<223> "n" may be a, c, g or t

<400> 7

```
Met Ser Ser Leu Thr Pro Pro Asn Ser Asn Gln Met Ser Ala Leu Asn
1               5                  10                  15

Asn His Phe Arg Phe Gly Leu Thr Thr Pro Glu Leu Glu Glu Phe Ala
            20                  25                  30

Pro Ala Leu Glu Ala Thr Leu Ala Ser Ser Glu Thr Val Glu Arg Leu
        35                  40                  45

Tyr Glu Arg Thr Ala Pro Glu Pro Pro Gln Arg Ser Trp Thr Ser Pro
    50                  55                  60

Thr Ala Asp Glu Asn Pro Leu Ser Ala Trp Tyr Val Thr Thr Ser Ile
65                  70                  75                  80

Ser Glu Thr Asp Glu Gly Pro Leu Ala Gly Arg Thr Val Ala Val Lys
                85                  90                  95

Asp Asn Val Ala Val Ala Gly Val Pro Met Met Asn Gly Ser Arg Thr
            100                 105                 110

Val Glu Gly Phe Thr Pro Arg Tyr Asp Ala Thr Val Val Arg Arg Leu
        115                 120                 125

Leu Asp Ala Gly Ala Thr Ile Thr Gly Lys Ala Val Cys Glu Asp Leu
    130                 135                 140

Cys Phe Ser Gly Ala Ser Phe Thr Ser His Pro Gln Pro Val Arg Asn
145                 150                 155                 160

Pro Trp Asp Glu Ser Arg Ile Thr Gly Gly Ser Ser Ser Gly Ser Gly
                165                 170                 175

Ala Leu Val Ala Ser Gly Gln Val Asp Met Ala Val Gly Gly Asp Gln
            180                 185                 190

Gly Gly Ser Ile Arg Ile Pro Ala Ala Phe Cys Gly Ile Val Gly His
        195                 200                 205

Lys Pro Thr His Gly Leu Val Pro Tyr Thr Gly Ala Phe Pro Ile Glu
    210                 215                 220

Arg Thr Ile Asp His Leu Gly Pro Met Thr Arg Thr Val Ser Asp Ala
225                 230                 235                 240
```

```
Ala Ala Met Leu Thr Val Leu Ala Gly Thr Asp Gly Leu Asp Pro Arg
                245                 250                 255

Gln Thr His Arg Ile Glu Pro Val Asp Tyr Leu Ala Ala Leu Ala Glu
            260                 265                 270

Pro Ala Ser Gly Leu Arg Val Gly Val Val Thr Glu Gly Phe Asp Thr
        275                 280                 285

Pro Val Ser Asp Ala Ala Val Asp Asn Ala Val Arg Thr Ala Ile Gly
    290                 295                 300

Val Leu Arg Ser Ala Gly Leu Thr Val Glu Glu Val Ser Ile Pro Trp
305                 310                 315                 320

His Leu Asp Ala Met Ala Val Trp Asn Val Ile Asp Arg Ala Asp Asp
                325                 330                 335

Glu Phe Glu Ala Phe Leu Leu Gln Val Leu Asp Glu Asn Ala Val Thr
            340                 345                 350

Ile Pro Glu Leu Gly Gln Val Arg Ala Gln Thr Pro Arg Ser Trp Cys
        355                 360                 365

Ser Pro Arg Thr Ala Pro Ala Arg Cys Thr Thr Pro Ser Asn Ala Ala
    370                 375                 380

Ala Cys Thr Thr Gly Ser Asn Thr Pro Thr Ser Arg Gly Lys Trp Arg
385                 390                 395                 400

Ser Cys Ala Ala Ala Ser Arg Ala Ser Thr Asn Thr Ser Arg Arg Arg
                405                 410                 415

Ser Pro Thr Pro Cys Arg Pro Cys Ala Gly Trp Thr Cys Ser Asn His
            420                 425                 430

Pro Gly Ser Arg Ser Arg Trp Thr Gly His Glu Arg Cys Gly Asn Ser
        435                 440                 445

Thr Ala Thr Cys Ser Thr Arg Arg Pro Arg Pro Arg Pro Ser Val Pro
    450                 455                 460

Ser
465
```